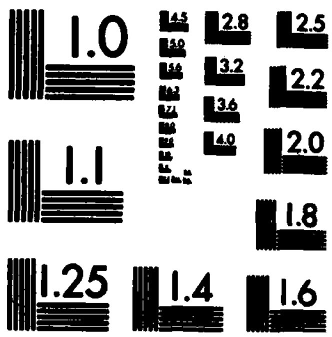
AD-A148 136 THE EM ALGORITHM FOR CENSORED DATA(U) NORTH CAROLINA UNIV AT CHAPEL HILL INST OF STATISTICS H SCHNEIDER ET AL. 1984 MIMEO SER-1557 AFOSR-TR-84-0939 1/1

UNCLASSIFIED F49620-82-C-0009 F/G 12/1 NL

END
FILMED
DTIC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0

FILE COPY

AD-A148 136

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

AD-A148136

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S): Mimeo Series #1557	5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR- 84-0939
6a. NAME OF PERFORMING ORGANIZATION: University of North Carolina; 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: Air Force Office of Scientific Research
6c. ADDRESS (City, State and ZIP Code): Department of Statistics, Chapel Hill NC 27514	7b. ADDRESS (City, State and ZIP Code): Directorate of Mathematical & Information Sciences, Bolling AFB DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: AFOSR; 8b. OFFICE SYMBOL (If applicable): NM	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: F 49620-82-C-0009
8c. ADDRESS (City, State and ZIP Code): Bolling AFB DC 20332-6448	10. SOURCE OF FUNDING NOS. PROGRAM ELEMENT NO.: 61102F; PROJECT NO.: 2304; TASK NO.: A5; WORK UNIT NO.:

11. TITLE (Include Security Classification)
THE EM ALGORITHM FOR CENSORED DATA

12. PERSONAL AUTHOR(S)
Lisa Weissfeld* and Helmut Schneider

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Technical	FROM ____ TO ____	1984	15

16. SUPPLEMENTARY NOTATION
*Dept of Biostatistics, University of North Carolina, Chapel Hill NC 27514.

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	Censored data; EM algorithm; linear regression; Kaplan-Meier estimator. expectation maximization

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Three methods for applying the ~~EM~~ algorithm to censored data are considered, the Buckley-James (1979), a proposed simpler nonparametric method and a normal model for censored data. A new estimator for the variance of y in the Buckley-James model is proposed and simulations comparing the three methods are described. To illustrate the use of these methods they are applied to the Stanford heart transplant data.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED X SAME AS RPT ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL: MAJ Brian W. Woodruff	22b. TELEPHONE NUMBER (Include Area Code): (202) 767-5027; 22c. OFFICE SYMBOL: NM

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

THE EM ALGORITHM FOR CENSORED DATA

H. Schneider* and L. Weissfeld**

*Free University of Berlin, Germany

**Department of Biostatistics, University of North Carolina at Chapel Hill

ABSTRACT

Three methods for applying the EM algorithm to censored data are considered, the Buckley-James (1979), a proposed simpler nonparametric method and a normal model for censored data. A new estimator for the variance of y in the Buckley-James model is proposed and simulations comparing the three methods are described. To illustrate the use of these methods they are applied to the Stanford heart transplant data.

Key Words: Censored data; EM algorithm; Linear regression; Kaplan-Meier estimator.

I. Introduction

There have been many methods proposed for handling regression problems in which the dependent variable may be censored. While several of these methods assume an underlying distributional form, that is, normal, Weibull or exponential, others are of a nonparametric nature and require minimal assumptions about the unspecified distribution.

One of these techniques, developed by Cox (1972), assumes that the hazard function $\lambda(y,x) = f(y,x)/(1-F(y,x))$ has the form

$$\lambda(y,x) = \lambda_o(y)e^{x\beta}$$

where $\lambda_o(y)$ is the underlying hazard. He then used conditional arguments to form a partial likelihood function (Cox, 1972, 1975), independent of $\lambda_o(y)$, for the estimation of β.

We will focus on three methods, which are based on the linear model

$$y = x^T\beta + \varepsilon.$$

Two of the methods considered here are nonparametric in that the distribution of y is unspecified while the third assumes that y is normally distributed. These methods rely on the expectation maximization (EM) algorithm, introduced by Dempster, Laird and Rubin (1977), which is broadly applicable for computing estimates from incomplete data and is used for the estimation of β and the variance of y by all three methods considered here. Although their presentation was based on a parametric model, the EM algorithm has been applied to nonparametric models (Buckley and James, 1979). In section 2 we discuss each of these methods and propose a new estimator of the variance of y for the Buckley-James method. In section 3 we present results from a simulation study which compares the performance of the three methods. Finally, in section 4, we present an example using data from the Stanford Heart transplant study.

2. Estimation Techniques for Linear Regression with Censored Data

We consider the linear model

$$Y_j = x_j^T\beta + \varepsilon_j \qquad j=1,\ldots, n \tag{1}$$

where the ε_j are independent and identically distributed with the distribution function F which has mean zero and finite variance. The covariate vector x_j^T is k dimensional with $x_{oj} \equiv 1$.

We assume here that the observed times are generated by random censorship. Let the life times $y_1,\ldots, y_n$ be *i.i.d.* with distribution function F as specified above and let the censoring times $C_1,\ldots, C_n$ be *i.i.d.* with distribution function G and further, under the assumption of random censorship, let $C_1,\ldots, C_n$ be independent of $y_1,\ldots, y_n$.

We observe

$$Z_j = \min(Y_j, C_j) \tag{2}$$

and

$$\delta_j = \begin{cases} 1 & \text{if} \quad Y_j \leq C_j \\ 0 & \text{if} \quad Y_j > C_j \end{cases} \tag{3}$$

with

$$n_u = \sum_{j=1}^{n} \delta_j \tag{4}.$$

For Type I censorship, a special case of random censorship, the C_j's are given constants.

Let X be the design matrix and define the vector $y^*(\beta)$ by

$$y_j^*(\beta) = \delta_j y_j + (1-\delta_j)\ E[Y_j \mid Y_j > C_j,\ x_j^T\beta]. \tag{5}$$

Then the EM estimate of β is the solution to

$$\hat{\beta} = (X^TX)^{-1}X^Ty^*\ (\hat{\beta}). \tag{6}$$

This solution requires an iterative procedure, since $E(Y_j \mid Y_j > C_j,\ x_j^T\beta)$ is unknown and has to be estimated.

The variance σ^2 can be estimated by

$$m\ \hat{\sigma}^2 = \left(y^*(\hat{\beta}) - X\hat{\beta}\right)^T \Delta \left(y^*(\hat{\beta}) - X\hat{\beta}\right) \tag{7}$$

where m is an appropriate constant and Δ is a diagonal matrix. It is not immediately evident what to substitute in (7) for $y_j^*(\hat{\beta})^2$ when y_j is censored. Aitkin (1981), who considers the case where F is normal, uses the maximum likelihood approach to derive the form

$$y_j^*(\beta)^2 = \delta_j y_j^2 + (1-\delta_j)\ E[Y_j^2 \mid Y_j > C_j, x_j^T\beta] \tag{8}$$

with $m = n$ and $\Delta = I$, the identity matrix. He suggests using the bias corrected estimate

$$\hat{\sigma}_a^2 = \hat{\sigma}^2 \frac{n_u}{n_u - k}. \tag{9}$$

Schmee and Hahn (1979), who also consider the normal model suggest

$$y_j^*(\beta)^2 = \delta_j y_j^2 + (1-\delta_j)\ E[Y_j \mid Y_j > C_j, x_j^T\beta]^2 \tag{10}$$

with $m = n-k$ and $\Delta = I$. But unfortunately, this approach results in an underestimation of σ^2 which becomes severe for moderate and heavy censoring.

Consequently, the estimate $\hat{\beta}$ suffers from poor estimation of σ^2. Since the results from the maximum likelihood estimator are better for moderate censoring we shall not report the results for the Schmee and Hahn estimator.

Buckley and James (1979) consider the nonparametric case where the underlying distribution F is unspecified. They suggest that the censored observations be ignored for the estimation of σ^2 and use

$$y_j^{*2} = \delta_j y_j^2, \; \Delta = \text{diag}(\delta_{ij}), \; m = n_u - k \tag{11}$$

but introduce the correction

$$\hat{\sigma}^2_{BJ} = \hat{\sigma}^2 - [\sum_j \delta_j (y_j - x_j^T\hat{\beta})/n_u]^2. \tag{12}$$

We propose a nonparametric estimator of σ analogous to that proposed by Aitken (1981) in equation (8). In fact, the simulation studies we performed suggest that this estimator of σ has a smaller bias and mean square error (MSE) than the estimator proposed by Buckley and James.

In order to apply the EM algorithm to equation (6) for the estimation of β we have to establish appropriate estimators for $E[Y_j \mid Y_j > C_j, x_j^T\beta]$. In the parametric case it is easy to find explicit expressions for this expectation and replace it by an appropriate estimate. In the normal model, where $F = \Phi$, we obtain (Aitkin (1981)),

$$E[Y_j \mid Y_j > C_j, x_j^T\beta] = x_j^T\beta + \sigma\, W((C_j - x_j^T\beta)/\sigma) \tag{13}$$

and

$$E[Y_j^2 \mid Y_j > C_j, x_j^T\beta] = (x_j^T\beta)^2 + \sigma^2 + \sigma(C_j + x_j^T\beta)\, W((C_j - x_j^T\beta)/\sigma). \tag{14}$$

where $W(u) = \phi(u) / (1 - \Phi(u))$ is the hazard rate of the normal distribution. In order to obtain an estimate for the latter two expectations we replace β and σ by their current estimates.

Buckley and James, who consider the case where the underlying distribution function F is unknown, estimate F using the product limit estimator (Kaplan and Meier, 1958),

$$1 - \hat{F}(e_j, \underset{\sim}{\beta}) = \prod_{\hat{e}_{(i)} \leq e} \left\{1 - \frac{d_{(i)}}{n_{(i)}}\right\}^{\delta_i}$$

where $\hat{e}_{(1)} < \hat{e}_{(2)} < \ldots < \hat{e}_{(n)}$ are the ordered values of the residuals $e_i(o,\hat{\beta}) = Z_i - x_i^T\hat{\beta}$, $n_{(i)}$ is the number at risk at $\hat{e}_{(i)}-$ and $d_{(i)}$ denotes the number of failures at $\hat{e}_{(i)}$. The Kaplan-Meier estimator is then used to estimate $E[Y_j \mid Y_j > C_j, x_j^T\beta]$ in the following manner. Uncensored observations are replaced by

$$\hat{E}[Y_j \mid Y_j > C_j, x_j^T\beta] = \hat{y}_j = x_j^T\hat{\beta} + \sum_{\kappa}^{u} w_{i\kappa}(\hat{\beta})(y_\kappa - x_\kappa^T\hat{\beta}) \tag{15}$$

where the sum is over the set of uncensored residuals and

$$w_{i\kappa}(\hat{\beta}) = \begin{cases} v_\kappa(\hat{\beta})/(1 - \hat{F}(C_i - x_i^T\hat{\beta})) \text{ when } e_i(o,\hat{\beta}) < e_\kappa(o,\beta) \\ 0 \qquad \text{otherwise} \end{cases} \tag{16}$$

where $v_\kappa(\hat{\beta})$ is the mass of the Kaplan-Meier estimator at the uncensored points. In this case the Kaplan-Meier estimator $\hat{F}(e,\underset{\sim}{\beta})$ assigns the remaining mass to the largest residual if it is censored.

We propose to estimate the variance σ^2 by applying the above reasoning to the computation of

$$\sum_{j=1}^{n} (y_j^*(\hat{\beta}) - x_j^T\hat{\beta})^2 = \sum_{j=1}^{n} (y_j^{*2}(\hat{\beta}) + (x_j^T\hat{\beta})^2 - 2y_j^*(\hat{\beta})x_j^T\hat{\beta}) \tag{17}$$

Replacing the $y_j^*(\hat{\beta})$ by their estimated expectations we have

$$n\hat{\sigma}^2 = \sum_{j=1}^{n} \hat{E}[y_j^{*2}(\beta)] + (x_j^T\hat{\beta})^2 - 2\hat{E}[y_j^*(\beta)]x_j^T\hat{\beta} \tag{18}$$

and from (15) we see that for censored observations

$$\hat{E}[y_j^{*2}(\beta)] = \hat{E}\left[(y_j^*(\beta) - x_j^T\beta)^2 + 2y_j^*(\beta)x_j^T\beta - (x_j^T\beta)^2\right]$$

$$= \sum_{\kappa}^{u} w_{j\kappa}(\hat{\beta})(y_\kappa - x_\kappa^T\hat{\beta})^2 + 2x_j^T\hat{\beta}\,\hat{E}(y_j^*(\beta)) - (x_j^T\hat{\beta})^2, \tag{19}$$

and

$$\hat{E}[y_j^{*2}(\beta)] + (x_j^T\beta)^2 - 2x_j^T\hat{\beta}\,\hat{E}[y_j^*(\beta)] = \sum_{\kappa}^{u} w_{j\kappa}(\hat{\beta})(y_\kappa - x_\kappa^T\hat{\beta})^2$$

substituting into equation (18) we have

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{j=1}^{n}\left\{\delta_j(y_j - x_j^T\hat{\beta})^2 + (1-\delta_j)\sum_{\kappa}^{u} w_{j\kappa}(\hat{\beta})\ (y_\kappa - x_\kappa^T\hat{\beta})^2\right\}.$$

This estimator, unlike the Buckley-James estimator, uses the information in both the censored and uncensored observations for the estimation of σ. To reduce bias we applied the bias correction

$$\hat{\sigma}^2_{new} = \hat{\sigma}^2 \cdot n_u/(n_u - k) \tag{20}$$

It is easy to see that the Buckley-James procedure becomes quite complicated, particularly when a large number of observations are tied. Therefore, we also consider a modified method which is very easy to implement. We propose estimating the expectation with

$$\hat{E}[Y_j \mid Y_j > C_j, x_j^T\beta] = x_j^T\beta + \sum_{\varepsilon(i)>e_j} \varepsilon(i)/\ell \tag{21}$$

where ℓ is the number of $\varepsilon(i) > c_j$. This is a simple average of all censored and uncensored residuals $\varepsilon_i = y_i^*(\hat{\beta}) - x_i^T\hat{\beta}$ exceeding $e_j = C_j - x_j^T\hat{\beta}$. This method, which is in spirit of the EM algorithm, replaces censored observations by their expectations and treats them as uncensored observations.

To estimate the covariance matrix of $\hat{\beta}$ in the normal model we calculate the inverse of the Fisher information matrix (Amemiya (1973)), which would be an approximation for the random censorship case. However, for the BJ estimator there are no theoretical results and Buckley and James suggest using the approximation

$$\mathrm{cov}(\hat{\beta}) = \hat{\sigma}^2 \{ (X-X^u)^T \Delta \ (X-X^u)^{-1} \} \tag{22}$$

where $\Delta = \mathrm{diag}(\delta_{ij})$ and X^u has elements $n_u^{-1} \sum_j \delta_j x_{ij}$. Equation (22) can also be used to approximate the covariance matrix of the simplified estimator.

3. Simulations

To gain some insight into the difference between the Buckley-James (BJ), simplified nonparametric (NP) and maximum likelihood (ML) methods simulation studies were carried out. In each case 1000 samples of size 50 were drawn with the covariates evenly spaced at $2i$, $i=1,\ldots, 50$. Independent life times (v_i) and censoring times (c_i) were generated for each covariate and $Z_i = \min(y_i, c_i)$ was modelled. The Beta distribution

$$f(x) = \frac{1}{B(p,q)} \frac{(x-a)^{p-1} (b-x)^{q-1}}{(b-a)^{p+q-1}}$$

was used to generate the censoring times because a wide variety of censoring patterns could be represented by varying the parameters, p and q, of the distribution. Symmetrical distributions are generated when p=q and various degrees and patterns of asymmetry can be generated by allowing p to be unequal to q. Each of the tables presents the average percent of censoring, means of the parameters, MSE of the parameters and the percentage of runs which did not converge. This percentage was computed using the total number of samples required to achieve 1000 convergent samples.

Table 1 presents the results for a normally distributed lifetime with equal censoring, that is, the percentage of censoring is the same at each design point. This case is of interest because the BJ estimator of σ is thought to be well-behaved under this assumption; however, we found that in general the BJ estimator had the largest bias and MSE of all estimators of σ considered. The best estimator of σ was the new estimator which is defined by equation (20). This estimator uses the information from the BJ results and makes full use of the censored data so that in general it has a smaller bias and MSE than the other estimators and, in particular, this estimator always behaves better than the BJ estimator. Both the BJ and NP estimators underestimated σ but the NP estimator was less biased than the BJ and was often less biased than the ML estimator of σ. The results for all three methods of estimating β_0 and β_1 were very similar with all methods being biased for the estimation of β_0. While the ML estimator performed better for the estimation of β_0 under heavy censoring (p=1, q=2), the BJ estimator was superior for the estimation of β_1 in this case. The percentage of nonconvergent samples was similar in most of the cases with the ML method having the most problems in the case of heavy censoring.

Table 2 presents the results for a normally distributed lifetime with increasing censoring. In general the simulations show that the ML estimate tends to overestimate σ while the BJ, the NP and the new estimators tend to underestimate σ. The NP estimator also tends to underestimate β_1; however, it has the smallest MSE in all the cases considered. Both the NP and ML estimators of σ tend to be better than the BJ estimator; however, the new estimator is generally superior to all the others in terms of bias and MSE. The results for the estimation of β_0 and β_1 tend to be very similar with the ML estimator doing very poorly in the case of heavy censoring (p=1, q=2).

Once again all the methods are biased for the estimation of β_o with the BJ performing the worst. As before, although the NP estimator of β_1 generally has a smaller MSE it is always biased downward.

Table 3 contains simulations where a beta distribution was used to generate both the lifetime and censoring distributions. The simulations were performed to explore the robustness of the ML method to departures from normality and to compare the behavior of the parametric and nonparametric methods in this setting. In general, the new estimator of σ had the smallest MSE and bias. In this instance the BJ estimator of σ behaved very poorly. For the symmetrical (p=q) lifetime distributions the ML estimator of β_1 tended to perform somewhat better than either of the other two, however, it was not consistently superior to the other estimators. The BJ estimator of β_o tended to be the least biased of the methods and often had the smallest MSE. For the case of the asymmetric lifetime distributions the BJ estimator of β_1 was the least biased of the three estimators but had the largest MSE, with the same results holding for β_o. Thus, the violation of the distributional assumption, although it does have some impact, did not seem to seriously affect the bias and MSE of the ML estimator.

TABLE 1

Equal Censoring for a Normally Distributed Lifetime

$\beta_0=0$, $\beta_1=0.2$

	censoring distribution beta p	q	% censoring	β_0	MSE	β_1	MSE (x 100)	σ	MSE	% nonconvergent samples
ML				0.117	13.95	0.199	0.400	10.48	2.92	1
NP	1	1	50	-0.639	14.19	0.200	0.397	9.79	2.18	5
BJ				-0.025	13.98	0.199	0.398	9.38(9.76)*	2.27(2.13)*	5
ML				0.013	13.24	0.199	0.364	10.50	2.86	1
NP	2	2	50	-0.939	13.42	0.199	0.350	9.61	2.17	6
BJ				-0.156	13.55	0.198	0.372	8.91(9.60)	2.87(2.08)	5
ML				-0.190	11.11	0.203	0.324	10.50	2.84	1
NP	3	3	50	-1.241	12.44	0.202	0.326	9.53	2.12	7
BJ				-0.379	12.29	0.202	0.349	8.61(9.53)	3.47(2.16)	6
ML				-0.032	8.60	0.201	0.247	10.21	1.38	0
NP	2	0.5	14	-0.113	8.67	0.202	0.253	9.63	1.27	3
BJ				-0.054	8.63	0.201	0.249	9.57(9.90)	1.31(1.14)	0.2
ML				0.322	11.73	0.195	0.344	10.28	1.79	0
NP	1	0.5	30	-0.001	11.67	0.196	0.346	9.74	1.45	4
BJ				0.283	11.75	0.195	0.345	9.60(9.88)	1.51(1.41)	1
ML				0.210	10.83	0.197	0.319	10.28	1.85	0
NP	2	1	28	-0.119	10.40	0.198	0.312	9.58	1.49	4
BJ				0.152	10.79	0.197	0.319	9.32(9.84)	1.74(1.47)	2
ML				0.186	9.81	0.197	0.291	10.22	1.52	0
NP	3	1	17	0.023	9.70	0.198	0.287	9.46	1.40	4
BJ				0.153	9.80	0.196	0.290	9.32(9.85)	1.57(1.30)	1
ML				-0.529	19.65	0.206	0.586	10.96	6.32	26
NP	1	2	71	-1.841	23.03	0.197	0.571	10.10	4.59	10
BJ				-0.195	22.14	0.196	0.615	9.13(9.46)	4.20(3.63)	12
ML				0.066	8.07	0.200	0.229	10.16	1.21	0
NP	9	0.5	1	0.037	8.00	0.201	0.228	9.76	1.03	0
BJ				0.055	8.03	0.201	0.228	9.76(9.92)	1.03(1.05)	0
ML				0.231	9.39	0.196	0.274	10.17	1.24	0
NP	9	1	2	0.191	9.26	0.197	0.271	9.63	1.09	1
BJ				0.203	9.31	0.196	0.273	9.61(9.90)	1.10(1.04)	0
ML				-0.178	8.49	0.203	0.238	10.21	1.25	0
NP	9	2	6	-0.264	8.28	0.203	0.236	9.37	1.32	2
BJ				-0.238	8.45	0.203	0.236	9.31(9.87)	1.39(1.03)	0.1

* in brackets new estimator for σ

TABLE 2

Increasing Censoring for a Normally Distributed Lifetime

$\beta_o=0$, $\beta_1=0.2$

	censoring distribution beta p	q	% censoring	β_o	MSE	β_1	MSE (x100)	σ	MSE	% nonconvergent samples
ML				-0.025	16.6	0.202	0.592	10.81	5.25	2
NP	1	1	66	-0.960	17.2	0.190	0.582	9.65	3.42	8
BJ				-0.220	17.4	0.200	0.615	9.00(9.52)*	3.89(3.10)*	9
ML				-0.200	15.7	0.202	0.573	11.01	6.57	4
NP	2	2	69	-1.457	16.9	0.182	0.544	9.75	3.82	8
BJ				-0.648	16.9	0.200	0.622	8.63(9.22)	4.84(3.90)	10
ML				-0.391	14.1	0.205	0.590	11.19	7.49	6
NP	3	3	72	-2.090	16.4	0.184	0.522	9.79	3.96	12
BJ				-1.079	15.7	0.205	0.658	8.35(9.06)	5.83(3.86)	12
ML				0.003	9.1	0.201	0.286	10.32	1.94	0
NP	2	0.5	30	-0.155	9.1	0.198	0.284	9.33	1.77	4
BJ				-0.109	9.2	0.201	0.292	9.02(9.77)	2.22(1.50)	2
ML				0.292	12.9	0.195	0.429	10.44	2.64	0
NP	1	0.5	46	-0.114	12.5	0.191	0.413	9.46	2.03	5
BJ				0.173	12.7	0.194	0.421	9.12(9.75)	2.41(1.91)	4
ML				0.124	11.8	0.200	0.392	10.50	2.91	0
NP	2	1	48	-0.365	11.3	0.192	0.372	9.49	2.15	4
BJ				-0.059	11.8	0.199	0.398	8.88(9.67)	3.04(2.10)	6
ML				0.177	10.5	0.198	0.346	10.38	2.21	0
NP	3	1	37	-0.080	10.1	0.192	0.330	9.34	1.91	5
BJ				0.038	10.4	0.197	0.346	8.84(9.70)	2.67(1.67)	3
ML				0.202	>100	0.225	20.40	14.00	>100	40
NP	1	2	81	-3.071	31.5	0.179	0.82	10.14	9.58	40
BJ				-1.498	29.6	0.200	1.061	8.84(9.19)	8.56(6.32)	35
ML				0.047	8.1	0.201	0.234	10.20	1.38	0
NP	9	0.5	9	0.047	8.1	0.199	0.234	9.18	1.59	3
BJ				0.013	8.0	0.200	0.237	9.08(9.81)	1.75(1.15)	1
ML				0.228	9.5	0.196	0.283	10.24	1.52	0
NP	9	1	14	0.156	9.3	0.195	0.280	9.11	1.75	4
BJ				0.156	9.4	0.196	0.287	8.90(9.79)	2.16(1.23)	2
ML				-0.173	8.9	0.203	0.275	10.31	1.72	0
NP	9	2	25	-0.286	3.8	0.198	0.267	9.07	2.00	3
BJ				-0.310	8.9	0.203	0.282	8.63(9.70)	2.95(1.41)	2

* in brackets new estimator for σ

TABLE 3

Increasing Censoring for a Beta Distributed Lifetime

β_1=0.2

	lifetime distribution beta p	q	censoring distribution beta p	q	% censoring	bias β_o	MSE	β_1	MSE (x100)	Bias σ	MSE	% noncon-vergent sample
ML						-1.08	55.68	0.199	1.596	1.09	16.17	3
NP	0.5	0.5	1	1	64	-4.88	63.43	0.187	1.253	-3.00	16.49	11
BJ						-1.94	52.91	0.204	1.927	-8.59(-1.73)*	79.84(10.06)*	12
ML						1.83	54.91	0.198	1.490	4.04	23.18	0
NP	0.5	0.5	9	0.5	31	-0.59	43.41	0.194	1.616	-2.15	8.13	3
BJ						-0.42	43.24	0.204	1.767	-4.99(-.11)	27.23(3.84)	1
ML						0.25	16.61	0.202	0.502	0.80	2.78	0
NP	2	2	9	0.5	15	0.06	16.44	0.197	0.506	-1.24	2.66	2
BJ						0.08	16.40	0.200	0.524	-1.82(-.14)	3.65(1.49)	1
ML						0.23	19.5	0.200	0.595	0.87	4.06	0
NP	2	2	3	1	39	-0.51	17.6	0.193	0.538	-1.06	3.01	3
BJ						-0.24	18.8	0.202	0.622	-2.26(-.29)	6.70(2.15)	4
ML						0.74	35.7	0.191	1.043	1.69	9.01	0
NP	1	1	3	1	42	-0.93	29.5	0.186	0.931	-1.62	5.97	6
BJ						-0.39	31.1	0.198	1.105	-4.23(-.59)	20.2(3.63)	6
ML						1.52	58.11	0.183	1.524	2.92	18.34	0
NP	0.5	0.5	3	1	55	-1.68	43.68	0.184	1.404	-2.17	9.59	7
BJ						-0.77	45.56	0.201	1.761	-7.22(-.85)	55.71(5.47)	6
ML						0.44	37.20	0.193	1.102	1.50	9.26	0
NP	1	1	2	1	50	-1.64	31.64	0.186	0.949	-1.60	6.25	9
BJ						-0.62	32.54	0.199	1.138	-4.51(-.79)	23.18(4.43)	7
ML						-0.27	7.22	0.201	0.200	-1.11	3.89	0
NP	0.5	3	1	1	31	-0.30	7.16	0.199	0.206	-2.11	6.67	1
BJ						-0.05	7.57	0.200	0.213	-2.38(-.51)	7.98(4.21)	0.3
ML						-0.13	16.88	0.200	0.539	0.15	3.77	0
NP	2	3	1	1	57	-1.25	15.98	0.197	0.471	1.18	3.79	4
BJ						-0.28	16.20	0.202	0.525	-2.02(-.45)	6.37(3.60)	6
ML						0.00	6.92	0.198	0.196	-0.39	2.28	0
NP	0.5	3	3	1	7	0.02	7.07	0.199	0.203	-1.39	3.54	1
BJ						0.04	7.12	0.200	0.207	-1.49(-.24)	3.86(2.49)	0.4
ML						0.10	12.64	0.198	0.392	0.19	2.08	0
NP	2	3	3	1	27	-0.17	11.97	0.197	0.374	-1.25	2.85	2
BJ						0.00	12.62	0.200	0.402	-1.77(-.25)	4.31(1.87)	1
ML						0.12	13.26	0.206	0.433	1.23	3.87	0
NP	3	2	9	0.5	24	-0.07	12.44	0.195	0.427	-0.70	1.86	1
BJ						-0.10	12.70	0.201	0.455	-1.30(-.05)	2.96(1.46)	1

* in brackets new estimator for σ

4. Stanford Heart Transplant Example

In order to further explore the relationship between the BJ, ML and NP estimators and to see how our proposed estimator of σ for the BJ method performed in a set of data, we analyzed the Stanford heart transplant data given in Miller and Halperin(1982). We modelled survival time (log 10) as a function of age for survival times greater than 10 days. These results, given in table 4, follow the pattern of the previous simulations with the new estimator (20) giving a larger estimate of σ than the BJ estimator. The ML estimate of the slope β is almost identical with the BJ estimate. The NP estimator results in a somewhat smaller estimate of β_0 and β_1 than the ML and BJ estimators.

	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\sigma}$	
MLE	3.826	-.02453	.890	
BJ	3.745	-.02434	.655	(.761)*
NP	3.645	-.0229	.751	

Table 4: Estimators for the heart transplant data.
*new estimator (20) of σ

5. Conclusions

Typically, all three methods for linear regression with censored data were very similar with no particular estimator of slope appearing superior. However, we did find that the new estimator of σ enhanced the BJ procedure and would cause any tests of significance for this model to be less anti-conservative so that we would recommend the use of this new estimator. We also found that the normal model tended to be fairly robust against departures from normality and in general performed quite well.

Helmut Schneider's work was supported by the Deutsche Forschungsgemeinschaft and the Air Force Office of Scientific Research.

Lisa Weissfeld's work was partially supported by National Heart, Lung and Blood Institute contract N01-HV-12243-L.

REFERENCES

AITKEN, M. (1981). A note on the regression analysis of censored data. Technometrics, 23, 161-163.

AMEMIYA, T. (1973). Regression analysis when the dependent variable is truncated normal. Econometrica, 41, 997-1016.

BUCKLEY, J. and JAMES, J. (1979). Linear regression with censored data. Biometrika, 66, 429-436.

COX, D.R. (1975). Partial likelihood. Biometrika, 62, 269-76.

COX, D.R. (1972). Regression models and life tables (with discussion). J. Roy. Statist. Soc., B34, 187-202.

DEMPSTER, A.P., LAIRD, N.M. and RUBIN, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). J. Roy. Statist. Soc., B39, 1-22.

MILLER, R. and HALPERIN, J. (1982). Regression with censored data. Biometrika, 69, 521-531.

SCHMEE, J. and HAHN, G.J. (1979). A simple method for regression analysis with censored data. Technometrics, 21, 417-432.

END

FILMED

12-84

DTIC